PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, 15/12, 15/62, 5/10, C07K 14/485, A01H 15/00

A1

(11) International Publication Number: WO 98/21348

(43) International Publication Date: 22 May 1998 (22.05.98)

(21) International Application Number: PCT/US97/20603

(22) International Filing Date: 12 November 1997 (12.11.97)

(30) Priority Data:
08/747,246 12 November 1996 (12.11.96) US

(71) Applicant: BATTELLE MEMORIAL INSTITUTE [US/US]; Pacific Northwest Division, Intellectual Property Services, P.O. Box 999, Richland, WA 99352 (US).

(72) Inventors: HOOKER, Brian, S.; 2525 W. Grand Ronde #14, Kennewick, WA 99336 (US). DAI, Ziyu; Apartment D-5, 1621 George Washington Way, Richland, WA 99352 (US). GAO, Jianwei; 506 Snyder Road, Richland, WA 99352 (US). KINGSLEY, Mark, T.; 531 North Reed Street, Kennewick, WA 99336 (US). WELLER, Richard, E.; 50 Herlou Place, Selah, WA 98942 (US).

(74) Agent: ZIMMERMAN, Paul, W.; Battelle Memorial Institute, Pacific Northwest Division, Intellectual Property Services, (K1-53), P.O. Box 999, Richland, WA 99352 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD OF PRODUCING HUMAN GROWTH FACTORS FROM WHOLE PLANTS OR PLANT CELL CULTURES

(57) Abstract

The production of hEGF is achieved in both whole plants and plant cell culture wherein the hEGF has a length of at least 200 amino acids. For epidermal growth factor this would comprise at least a tetramer of EGF units. Effectiveness or production of the translation process has been increased according to the present invention by (1) cloning of pre-pro-EGF cDNA of approximately 4.5 kb into both whole plants and cell culture to increase overall titers of active hEGF; (2) synthesizing cDNA and transforming plants and cell culture for production of an oligomeric polypeptide consisting of repeated hEGF domains; and (3) increasing the overall size of the gene to be expressed with a fusion construct encoding hEGF linked to a protein that is efficiently produced in plant systems. As needed, synthetic cDNA includes plant-specific proteolytic cleavage sites between EGF repeats to facilitate correct processing *in planta*. Appropriate proteolytic cleavage sites upstream and downstream of hEGF are added if needed to obtain final product. In whole plants, use of a regulatory element confers hEGF production characteristics into traditionally non-saleable portions of crop plants, such as the leafy tops of potatoes. Use of potato tops under post-harvest conditions, results in overexpression production of hEGF in non-saleable plant portions towards the end of the harvesting season, without affecting crop quality.

BNSDOCID: <WO_______9821348A1_I_>

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

BNSDOCID: <WO_________9821348A1_I_>

METHOD OF PRODUCING HUMAN GROWTH FACTORS FROM WHOLE PLANTS OR PLANT CELL CULTURES

FIELD OF THE INVENTION

The present invention relates generally to a method for producing human growth factors from whole plants or plant cell culture. More specifically, the invention relates to producing a human growth factor from a plant cell encoded to produce the human growth factor with a length of at least 200 amino acids from transgenic plant cells.

BACKGROUND OF THE INVENTION

Growth factors and monoclonal antibodies (Mabs) are diverse yet highly specialized types of proteins having research and commercial applications in areas of therapeutics and diagnostics.

Therapeutic uses of human epidermal growth factor (hEGF) include treatment of soft tissue wounds (U.S. 5,218,093, 1993), specifically including skin and eye injuries as well as corneal and stomach ulcers (Frost and Sullivan 1996, 1994). In addition, several hEGF-bearing fusion constructs have been considered and/or tested, including mitotoxins for treatment of restenosis (Frost and Sullivan, 1994) and radioconjugates for a variety of anti-neoplastic therapies (Grieg et al., 1988).

Current production techniques for these proteins such as hybridoma and other types of mammalian cell culture methods (Köhler and Milsten, 1975) are generally slow, labor intensive, and consequently, expensive. In addition, current production techniques are difficult to validate due to the pathogenic and oncogenic potential of cultivated mammalian tissue.

Multimers of from 2 to 7 EGF units each having 53 amino acid residues have been produced from bacterial hosts, eg E. coli, Streptomyces and Bacillus, fungal hosts, eg Saccharomyces, Pichia and Aspergillus, insect cell host, and

mammalian cell hosts, eg CHO cells and COS cells. (U.S. Patent No. 5218093, 1993). hEGF production in *Staphylococcus aureus* (U.S. Patent No. 5004686, 1991) is by a fusion construct encoding hEGF linked to a protein. Synthesis methods using transgenic bacterial strains have problems such as faulty antibody gene expression, protein folding difficulties, inability to glycosylate proteins, and relegation of foreign peptides to insoluble material accumulated in inclusion bodies.

Transgenic plants can be used for the production of high value, medicinally important proteins, for example, production of Mabs (Hiatt et al., 1989; Düring et al., 1990; Benvenuto et al. 1991, Firek et al. 1993, Gao et al. 1993), human growth hormone (Kay et al. 1986) and human serum albumin (Sijmons et al. 1990). Transformed cells synthesize, secrete, and accumulate functional antibodies including single (Benvenuto et al. 1991) and double (Düring et al. 1990, Hiatt et al. 1991) domain immunoglobulins. However, it is noted that none of these authors investigated production of any human growth factor from transgenic plants.

Plant cell culture media are well-defined and inexpensive compared to mammalian cell culture media. Further, plant cell products, unlike mammalian-derived protein formulations, are generally assumed as neither pathogenic nor oncogenic to humans (Crawford, 1995). Also, when compared to similar production in transgenic bacterial strains (Attaai and Shuler 1987), plant tissue culture methods showed greater stability of foreign gene expression, even without use of selection pressure (Gao et al. 1991). One author, Higo et al. (1993) produced a human growth factor, specifically hEGF in transgenic tobacco with cDNA fragment size of 180 bp. Unsatisfactory foreign peptide levels of 20 to 60 pg/mg (ppb) total soluble leaf protein were obtained. This is despite the fact that plant progeny appeared to produce high levels of hEGF mRNA. Exact reasons for low observed levels of hEGF production are unclear. However, no signal peptide was encoded upstream of hEGF cDNA which could cause the foreign protein to be

relegated to the cytosol. Within this cell fraction, hEGF suffers proteolytic attack, especially considering the relatively small size (53 amino acids) or the peptide.

Although advantages have been observed for deriving proteins including EGF from plants, no transgenic plant cell culture process has been commercially developed for production of human growth factor. The lack of commercial exploitation of plant derived proteins is due in part to existing technological hurdles as observed by Higo et al. In addition, Ma et al., 1995 reported Mab titers of up to 500 μg/g (ppm) fresh weight of plant material (or 300 mg/L on a cell culture basis) whereas comparable mammalian cell processes are reported to attain levels of 1-2 g/L and higher (Rosenberg, personal communication, 1995). Implementation of alternative production systems to mammalian and bacterial culture, such as plant cellular techniques, has been further limited by non-technological factors, such as industry and regulatory acceptance (Simonsen and McGrogan, 1994) because of the investment made in developing and validating the more established non-plant methods.

Accordingly, there is a continuing need for plant based production of human growth factors.

<u>Background References</u>

1. U.S. Patent No. 5218093, 1993.

2. Frost & Sullivan. Emerging wound management technologies: wound healing/wound closure growth. February 1996.

3. Frost & Sullivan. World prescription dermatology pharmaceuticals markets. June 1994.

4. Grieg R, Dunnington D, Murthy U, Anzano M. 1988. Growth factors as novel therapeutic targets in neoplastic disease. *Cancer Surveys* 7(4):653-674.

5. Köhler G, Milsten C. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*. 256:495.

6. U.S. Patent No. 5004686, 1991.

7. Hiatt A, Cafferkey R, Bowdish K. 1989. Production of antibodies of transgenic plants. *Nature.* 342:76-78.

8. Düring K, Hippe S, Kreuzaler F, Shell J. 1990. Synthesis and self-assembly of a functional monoclonal antibody in transgenic *Nicotiana tabacum*. *J Plant Mol Biol* 15:281-293.

9. Benvenuto E, Ordas RJ, Tavazza R, Ancora G, Biocca S, Cattaneo A, Galeffi P. 1991. 'Phytoantibodies': a general vector for the expression of immunoglobulin domains in transgenic plants. *Plant Mol Biol* 17:865-874.

10. Firek S, Draper J, Owen MRL, Gandecha A, Cockburn B, Whitelam GC. 1993. Secretion of a functional single-chain Fv protein in transgenic tobacco plants and cell suspension cultures. *Plant Mol Biol* 23:861-870.

11. Gao J, Konzek RL, Linzmeier M, Buckley KB, Magnuson NS, Reeves R, An G, Lee JM. Production of monoclonal antibodies in plant cell culture. Presented at the 1993 ACS Fall Symposium, Denver.

12. Kay R, Chan A, Dayly M, McPherson J. 1987. Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236:1299-1302.

13. Sijmons PC, Dekker BMM, Schrammeieijer B, Verwoerd TC, van den Elzen JM, Hoekema A. 1990. Production of correctly processed human serum albumin in transgenic plants. *Bio/Technology* 8:217-221.

14. Crawford M. 1995. Therapeutic protein production using plant cell culture. Report by Lasure and Crawford, Inc.

15. Attaai MM, Shuler ML. 1987. A mathematical model for predicting of plasmid copy number and genetic stability in Escherichia coli. *Biotechnol Bioeng* 30:389-397.

16. Gao J, Lee JM, An G. 1991. The stability of foreign protein production in genetically modified plant cells. *Plant Cell Reports* 10:533-536.

17. Higo K, Saito Y, Higo H. 1993. Expression of a chemically synthesized gene for human epidermal growth factor under the control of cauliflower mosaic virus 35S promoter in transgenic tobacco. *Biosci Biotechnol Biochem* 57:1477-1481.

18. Ma JKC, Hiatt A, Hein M, Vine N, Wang F, Stabila P, van Dolleweerd C, Mostov K, Lehner T. 1995. Generation and Assembly of Secretory Antibodies in Plants. *Science* **268:716-719.**

19. Simonsen CC, McGrogan M. 1994. The molecular biology of production cell lines. *Biologicals* 22:85-94.

SUMMARY OF THE INVENTION

Despite the hurdles in technology development and commercialization, economic analysis indicates that regulatory costs associated with plant cell culture may reduce by as much as $70,000 per batch as compared to analogous mammalian cell processes (Crawford, 1995). In addition, direct production costs for whole plant processes at equal protein production rates appear to be two to four orders-of-magnitude lower than comparable mammalian cell processes (Agracetus 1995). Additionally, as plant cell titers increase, this type of production becomes even more capital cost-effective.

It is, therefore, an object of the present invention to provide whole plant and plant cell culture derived human growth factors at higher overall concentrations and production rates, comparable to mammalian host cell systems.

It is a further object of the present invention to synthesize specific human growth factors.

It is another object of the present invention to increase production rates and concentrations by increasing protein stability through the use of fusion constructs.

It is a further object of the present invention to use *Nicotiana tabacum* (tobacco) and *Solanum tuberosum* (potato) whole plants and highly synchronous suspensions.

According to the present invention, the production of human growth factors is achieved in whole plants or plant cell culture wherein the human growth factor is produced with a length of at least 200 amino acids. For epidermal growth factor this would comprise at least a tetramer of EGF units.

Modifying chimeric cDNA and subcloning into a plant expression vector are done using standard molecular cloning procedures (Ausubel et al. 1992) and splicing PCR techniques (Marks et al. 1992).

Effectiveness or production of the translation process has been increased according to the present invention by (1) cloning of pre-pro-EGF cDNA of approximately 4.5 kb into both whole plants and cell culture to increase overall titers of active hEGF, (2) synthesizing cDNA and transforming plants and cell culture for production of an oligomeric polypeptide consisting of repeated hEGF domains, and (3) increasing the overall size of the gene to be expressed with a fusion construct encoding hEGF linked to a protein that is efficiently produced in plant systems. As needed, synthetic cDNA includes plant-specific proteolytic cleavage sites between EGF repeats to facilitate correct processing *in planta*. Appropriate proteolytic cleavage sites upstream and downstream of hEGF are added if needed to obtain final product.

The subject matter of the present invention is particularly pointed out and distinctly claimed in the concluding portion of this specification. However, both the organization and method of operation, together with further advantages and objects thereof, may best be understood by reference to the following description taken in connection with accompanying drawings wherein like reference characters refer to like elements.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 provides the size of EGF precursor (pre-pro-EGF) relative to correctly processed EGF.

FIG. 2 depicts schematically the construction of pZD203, a vector used to modify the restriction sites on pre-pro-EGF to develop cDNA suitable for cloning into the plant expression vector pGA643.

FIG. 3 depicts schematically the construction of pZD204, the plant expression vector carrying pre-pro-EGF.

FIG. 4 shows EGF levels seen in individual calli resulting from positive transformation and antibiotic selection. EGF concentrations were determined using enzyme-linked immunosorbent assay and are based on a 30 KD protein size.

DESCRIPTION OF THE PREFERRED EMBODIMENT(s)

The present invention is a method for production of human growth factors using whole plants as well as plant cell suspensions transformed with appropriately constructed vector plasmids, wherein the human growth factor is produced with a length of at least 200 amino acids. More specifically, the method of the present invention is stable expression of human growth factors of interest as direct therapeutics, targeted delivery systems and research reagents. Human growth factors produced include human epidermal growth factor (hEGF), transforming growth factor (TGF), vascular endothelial growth factor (VEGF), platelet-derived growth factor (PDGF), fibroblast growth factor (FGF), tumor necrosis factor (TNF), heparin-binding epidermal growth factor (HBEGF), insulin-like growth factor (ILGF), platelet-derived endothelial cell growth factor (PDECGF), platelet-derived angiogenesis factor (PDAF), and bone-and-cartilage inducing growth factor (BCIF).

Any plant from the plant kingdom may be utilized. Specific types of plants that are amenable to the transformation steps listed herein include, but are not limited to monocotyledonous, dicotyledonous, and tuberous plants. Preferred species include but are not limited to *Nicotiana tabacum* (tobacco), *Solanum tuberosum* (potato), *Glycine max* (soybean), and *Zea mays* (corn).

The method of the present invention, a method of producing human growth factors from plant cells, has the steps of:

(a) obtaining a positive transformant of the plant cells, the positive transformant carrying genetic material encoding the production of a human growth factor with a length of at least 200 amino acids;

(b) cultivating the positive transformant; and

(c) obtaining the human growth factors.

The step of obtaining may be as simple as purchasing or more complex actual making by well known methods, for example direct particle bombardment as described in Gene Transfer by Particle Bombardment, Klein TM, Knowlton S, Arentzen R, Plant Tissue Culture Manual, D1, pp 1-12, 1991, Kluwer Academic Publishers, or by Agrobaterium mediated transformation as described in Hoekema et al. 1985 (Hoekema KM, Hirsch PR, Hooykaaf PJJ, Schliperoort RA, 1985, Nononcogenic Plant Vectors for Use in the Agrobacterium Binary System, Plant Molecular Biology, Vol. 5, 85-89), and further described herein.

The step of cultivating involves either whole plant cultivating or tissue cultivating by any of well known cultivating methods.

The step of obtaining is by well known separation purification steps, for example ultrafiltration, affinity chromatography, and/or electrophoresis.

An Agrobacterium mediated transformation method of the present invention has the steps of:

(a) modifying chimeric cDNA encoding a specific growth factor for subcloning into a plant expression vector

(b) subcloning the chimeric cDNA into the plant expression vector;

(c) transferring the plant expression vector containing transgenic plant cells to an agrobacterium;

(d) co-cultivating a portion of the transgenic plant cells (suspension culture or leaf disks) with the agrobacterium;

(e) selecting positive transformants from the co-cultivated culture on an antibiotic selective media;

(f) permitting growth of the transgenic plant cells in whole plants or suspensions; and

(g) extracting a liquid containing the human growth factor;

wherein the improvement comprises:

said human growth factor having a length of at least 200 amino acids.

Modifying chimeric cDNA and subcloning into a plant expression vector are done using standard molecular cloning procedures (Ausubel et al. 1992) and splicing PCR techniques (Marks et al. 1992). More specifically, modifying chimeric cDNA, has the steps of:

(a) adding a transcription promoter to the upstream or 5' end of the chimeric cDNA; and

(b) adding a transcription terminator to the downstream or 3' end of the chimeric cDNA. The transcription promoter and the transcription terminator are regulatory elements.

Further, an additional regulatory element encoding a signal peptide may be added between the transcription promoter and the 5' end of the chimeric cDNA in order to relegate the product human growth factor to a specific cellular organelle. In addition, other regulatory elements may be added either between the promoter and the additional regulatory element encoding the signal peptide or at the 3' end of the chimeric cDNA to obtain greater mRNA stability between transcription and translation events.

In either whole plants or cell cultures, to enhance expression of the chimeric gene (hEGF), the present invention further includes manipulation of a 35S promoter by duplication of the upstream region (-343 to -90 bp) of the CaMV 35S promoter to increase transcription activity, as well as use of TSC29 and TSC40 promoters. These promoters and their transcription activity have been reported by Gao et al. 1994, and Dai et al. 1995.

In whole plants, transcription promoters may include the upstream enhancer (nucleotides -343 to -90 relative to the transcription start site) of the CaMV 35S promoter (Benfey et al. 1989) or the chlorophyll a/b binding protein (*cab1*) promoter (Ha and An 1988). Use of these types of regulatory elements confers human growth factor production characteristics into traditionally non-salable portions of crop plants, such as the leafy tops of potatoes. Use of potato tops, for example, under post-harvest conditions, results in overexpression and production

of human growth factor in non-salable plant portions towards the end of the harvesting season, without affecting crop quality.

Transferring the plant expression vector into the agrobacterium is completed using the freeze-thaw method (An 1987). For monocotyledonous species, super-binary vectors, such as pTOK233 and pSB131, are used to achieve high transformation frequency (Ishida et al. 1996). Remaining cocultivation, selection, growth, and extraction steps (d through g) have been described by Magnusen et al. (1996), and are well known in the art of plant molecular biology.

Many human growth factors possess relatively short lengths of between 50 and 100 amino acids. For example, hEGF has a length of 53 amino acids. Accordingly, obtaining a larger construct of at least 200 amino acids requires either (1) cloning the larger precursor cDNA, (2) synthesizing a concatemer consisting of multiple gene copies encoding the growth factor, or (3) increasing the overall size of a gene to be expressed using a fusion construct encoding a growth factor linked to a protein that is efficiently produced in plant systems.

An example of obtaining a larger precursor to increase the overall protein size is the cDNA encoding pre-pro-EGF. This particular gene, at approximately 4.5 kb, encodes a 1207 amino acid protein that, *in vivo*, is proteolytically cleaved to yield 53 amino acid EGF. In plant systems, this larger protein will provide additional stability against proteolytic degradation.

Synthesizing the cDNA concatemer is preferably done by ligating multiple gene copies using peptide linkers to obtain a processed protein length of at least 200 amino acids. The multiple gene copies are preferably an oligomeric polypeptide having of repeated growth factor cDNA domains. Peptide linkers may be used that are (1) proteolytically cleaved *in planta*, (2) proteolytically cleaved in a separate enzymatic treatment step, or (3) resistant to proteolytic cleavage. Peptide linkers that are proteolytically cleaved by serine proteases *in planta* preferably possess the amino acid sequence Arg-Asn. This sequence already exists when EGF is concatemerized since the C-terminal amino acid is arginine and the N-terminal amino acid is asparagine. To achieve *in planta* cleavage, the processed

protein is targeted either to the cell cytosol (no signal peptide) or vacuole (phytohemagglutinin signal peptide [Chrispeels et al.1991]). To achieve proteolytic cleavage in a separate enzymatic treatment step, the same amino acid sequence is preferably used (Arg-Asn) and the growth factor concatemer is either targeted to the chloroplast (pea photosystem II signal peptide) or secreted (PR-II signal peptide) to limit proteolytic degradation. To achieve resistance to proteolytic cleavage, linkers would preferably possess the amino acid sequence Arg-Pro. This sequence is resistant to serine proteases. Specifically for EGF, linkage would preferably be achieved by synthesizing cDNA encoding a single proline unit between growth factor monomers cDNA.

Increasing the overall size of a gene may be done by ligating EGF with cDNA encoding a protective protein to protect from proteolytic cleavage, thereby forming a fusion construct. Protective proteins include but are not limited to streptococcal protein G or -galactosidase, that have both been shown to inhibit proteolysis when attached to the C-terminus of other foreign proteins (Hellebust et al. 1989). Gene size could also be increased by ligating EGF with cDNA encoding another protective protein of commercial interest that processes well in plant-based systems. Protective proteins further include human serum albumin (Sijmons et al. 1990) and phytase (Verwoerd et al. 1995).

At least one genetic regulatory element may be included in the cDNA encoding the transcription of specific growth factors. Regulatory elements include transcription promoters or enhancers that increase the frequency of transcription events, leader sequences that increase the stability of mRNA prior to translation, and signal peptides that target proteins to specific organelles for posttranslational modifications and accumulation. Examples of transcription enhancers include but are not limited to the octapine synthase enhancer, a 16 bp palindrome (ACGTAAGCGCTTACGT) (Ellis et al. 1987) and the B-domain of the cauliflower mosaic virus 35S promoter (Kay et al. 1987). An example of a leader sequence includes but is not limited to alfalfa mosaic virus RNA4 leader sequence (Jobling and Gehrke 1987). Examples of signal peptides include but are not

limited to the tobacco PR-S signal peptide (Cornelissen et al. 1986) and the phytohemagglutinin signal peptide (Hunt and Chrispeels 1991).

Example 1

The bacteriophage λEGF116 (ATCC No. 59956) containing the gene encoding the full length polypeptide of human kidney pre-pro-EGF was obtained from ATCC. Pro-EGF (**FIG. 1**) is the 1207 amino acid precursor in which hEGF is flanked by polypeptide segments of 907 and 184 residues at its NH_2- and COOH-termini, respectively (Bell et al., 1986). The remainder of the 4.8 kb pre-pro-EGF gene encodes native signal peptides at both the NH_2- and COOH- termini of pro-EGF. The polypeptide contains a transmembrane (TM) binding region that facilitates proper cleavage in the endoplasmic reticulum.

The full length of cDNA was excised with Sma I, Hind III, and Eco RI restriction enzymes, as shown on **FIG. 2**, producing two separate fragments. These were sequentially ligated into compatible Sma I and Eco RI sites in pBluescript- creating the 7.5 kb plasmid pZD203. After proper orientation was confirmed, pre-pro-EGF cDNA was further excised with Xba I and Cla I restriction enzymes and ligated into compatible sites located between the CaMV 35S promoter and T_7 transcription terminator of binary vector pGA643, forming the 16 kb plasmid pZD204 (**FIG. 3**). This plasmid was directly transferred into *Agrobacterium tumefaciens* LBA4404 using the freeze-thaw method (An 1987). The transferred plasmid was introduced into tobacco whole plants (by leaf disks) and calli (by suspension culture) by co-cultivation with the Agrobacterium thereby producing transformants. Over 200 specific samples of transformants were taken from the co-cultivation and separately placed on kanamycin selective media. The co-cultivated transformants that grew were positive transformants. The positive transformants were screened under kanamycin selection pressure and preliminary ELISA results indicated the presence of hEGF in tobacco calli. Accumulation levels of hEGF in select transgenic calli are shown on a ng/g fresh weight basis in **FIG. 4**. The bars in FIG. 4 represent a random sample of the specific samples of

transformants. The highest level of accumulation at approximately 400 ng/(g fresh weight cells) (ppb) corresponds to a concentration of 4.1 ng/(mg total soluble protein) (ppm) (based on a measured total soluble protein level of approximately 98 mg/(g fresh weight cells)). The 4.1 ng/(mg total soluble protein) (ppm) corresponds to 4100 pg/(mg total soluble protein) (ppb) which is almost two orders-of-magnitude greater than the result of 60 pg/(mg total soluble protein) (ppb) reported by Higo et al. (1993).

Further ELISA and Northern blot analyses were used to detect high levels of foreign protein production and mRNA transcription, respectively. Western blot analysis, completed to determine protein size, showed that specific EGF bearing constructs of 30 KD were produced. This size corresponds to approximately 250 amino acids.

Closure

While a preferred embodiment of the present invention has been shown and described, it will be apparent to those skilled in the art that many changes and modifications may be made without departing from the invention in its broader aspects. The appended claims are therefore intended to cover all such changes and modifications as fall within the true spirit and scope of the invention.

References

1. Crawford M. 1995. Therapeutic protein production using plant cell culture. Report by Lasure and Crawford, Inc., Seattle, Washington.

2. Tuan J. 1995. Plant bioreactor systems program overview. Report by Agracetus, Inc., Middleton, Wisconsin.

3. Benfey PN, Ren L, Chua NH. 1989. The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. *EMBO J* 8:2195-2202.

4. Ha SB, An G. 1988. Identification of upstream regulator elements in the developmental expression of the Arabidopsis thaliana cab1 gene. *Proc Natl Acad Sci USA* 85:8017-8021.

5. Gao J, Kim SR, Chung YY, Lee JM, An G. 1994. Developmental and environmental regulation of ribosomal protein genes in tobacco. *Plant Mol Biol* 25:761-770.

6. Dai Z, Gao J, An G. 1996. Regulatory elements controlling developmental and environmental regulation of a ribosomal protein L34 in tobacco. *Plant Mol Biol* In press.

7. Ausubel FM, Brent R, Kingson RE, Moore DD, Seidman JG, Smith JA, Struhl K. 1992. Current protocols in molecular biology. John Wiley & Sons, New York.

8. Marks JD, Hoogenboom HR, Griffiths AD, Winter G. 1992. Molecular evolution of proteins on filamentous phage. Mimicking the strategy of the immune system. *J Biol Chem* 267:16007-16010.

9. An G. 1987. Binary Ti vectors for plant transformation and promoter analysis. *Meth Enzy* 153:292-305.

10. Ishida Y, Saito H, Ohta S, Hiei Y, Komari T, Kumashiro T. 1996. High efficiency transformation of maize (*Zea mays* L.) Mediated by *Agrobacterium tumefaciens*. *Nature Biotechnol* 14:745-750.

11. Magnuson NS, Linzmaier PM, Gao J, Reeves R, An G, Lee JM. 1996. Enhanced recovery of a secreted mammalian protein from suspension culture of genetically modified tobacco cells. *Prot Expr Purif* 7:220-228.

12. Chrispeels MJ, Dickinson CD, Tague BW, Hunt DC, von Schaewen A. 1991. Defining the vacuolar targeting signal of phytohemagglutinin. In: Plant Molecular Biology 2. Hermann RG, Larkins B, Eds. Plenum Press, New York. pp. 575-582.

13. Hellebust H, Uhlen M, Enfors SO. 1989. Effect of protein fusion on the stability of proteolytically sensitive sites in recombinant DNA proteins. *J Biotechnol* 12:275-284.

14. Sijmons PC, Dekker BMM, Schrammeieijer B, Verwoerd TC, van den Elzen JM, Hoekema A. 1990. Production of correctly processed human serum albumin in transgenic plants. *Bio/Technology* 8:217-221.

15. Verwoerd TC, van Paridon PA, van Ooyen AJJ, van Lent JWM, Hoekema A, Pen J. 1995. Stable accumulation of Aspergillus niger phytase in transgenic tobacco leaves. *Plant Physiol* 109:1199-1205.

16. Ellis JG, Llewellyn DJ, Walker JC, Dennis ES, Peacock WJ. 1987. The ocs element: a 16 base pair palindrome essential for activity of the octopine synthase enhancer. *EMBO J* 6:3203-3208.

17. Kay R, Chan A, Dayly M, McPherson J. 1987. Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236:1299-1302.

18. Jobling SA, Gehrke L. 1987. Enhanced transcription of chimaeric messenger RNAs containing a plant viral untranslated leader sequence. *Nature* 325:622-625.

19. Cornelissen BJC, Hooft van Huysduynen RAM, Bol JF. 1986. A tobacco mosaic virus-induced tobacco protein is homologous to the sweet-tasting protein thaumatin. *Nature* 321:531-532.

20. Hunt DC, Chrispeels MJ. 1991. The signal peptide of a vacuolar protein is necessary and sufficient for the efficient secretion of a cytosolic protein. *Plant Physiol* 96:18-25.

21. Bell GI, Fong NM, Stempien MM, Wormsted MA, Caput D, Ku, L, Urdea MS, Rall LB, Sanchez-Pescador R. 1986. Human epidermal growth factor precursor: cDNA sequence, expression *in vitro* and gene. *Nucleic Acids Res* 14:8427-8446.

22. Higo K, Saito Y, Higo H. 1993. Expression of a chemically synthesized gene for human epidermal growth factor under the control of cauliflower mosaic virus 35S promoter in transgenic tobacco. *Biosci Biotechnol Biochem* 57:1477-1481.

```
(1) GENERAL INFORMATION:
     (i)   APPLICANT:  Brian S. Hooker, et al
     (ii)  TITLE OF INVENTION: Method of Producing Human Growth
     Factors From Whole Plants or Plant Cell Cultures
     (iii)NUMBER OF SEQUENCES: 8
     (iv)  CORRESPONDENCE ADDRESS:
                (A)  ADDRESSEE:  Paul W. Zimmerman
                (B)  STREET:        P.O. Box 999
                (C)  CITY:       Richland
                (D)  STATE:         WA
                (E)  COUNTRY:       USA
                (F)  ZIP:           99352
     (v)   COMPUTER READABLE FORM:
                (A) MEDIUM TYPE: 3 1/2 Magnetic Disk
                (B) COMPUTER: IBM compatible
                (C) OPERATING SYSTEM: DOS
                (D) SOFTWARE: WORD97
     (vi)  CURRENT APPLICATION DATA:
                (A) APPLICATION NUMBER: 08/747,246
                (B) FILING DATE: 11-12-96
                (C) CLASSIFICATION: unknown
     (vii) PRIOR APPLICATION DATA:
                (A) APPLICATION NUMBER:  N/A
                (B) FILING DATE:
     (viii) ATTORNEY/AGENT INFORMATION:
                (A) NAME:  Paul W. Zimmerman
                (B) REGISTRATION NUMBER: 34,761
                (C) REFERENCE/DOCKET NUMBER: E-1519
     (ix)  TELECOMMUNICATION INFORMATION:
                (A) TELEPHONE: 509-375-2981
                (B) TELEFAX: 509-375-2592
                (C) TELEX:
```

(2) INFORMATION FOR SEQ ID NO: 1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4481bp
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: double strands
(D) TOPOLOGY: unknown
(ii) MOLECULE TYPE:
(A) DESCRIPTION: Sense orientation of complementary DNA for pro-EGF
(iii) HYPOTHETICAL:
(iv) ANTI-SENSE: 5'-AGT GAC TCA GTC GAG ... TTC TCA CTC GTC-3 end
(v) FRAGMENT TYPE: 4.5kb Smal/Hindlll double strands DNA fragment
(vi) ORIGINAL SOURCE:
(A) ORGANISM: kidney
(B) STRAIN: human
(C) INDIVIDUAL ISOLATE: GI Belle
(D) DEVELOPMENTAL STAGE: adult
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: fetal human liver library
(B) CLONE: lambda CH4A; lambda EMBL4; lambda GM1416
(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT:
(B) MAP POSITION:
(C) UNITS:
(ix) FEATURE:
(A) NAME/KEY: human epithelial growth factor cDNA
(B) LOCATION:
(C) IDENTIFICATION METHOD: cross-hybridization with mouse cDNA
(D) OTHER INFORMATION:
(x) PUBLICATION INFORMATION:
(A) AUTHORS:
(B) TITLE:
(C) JOURNAL:
(D) VOLUME:
(E) ISSUE:
(F) PAGES:
(G) DATE:
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO:
FROM ____(position) TO ____ (position)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
CCCGGGCCAT GCTCCAGCAA AATCAAGCTG TTTTCTTTTG AAAGTTCAAA CTCATCAAGA TT    62

ATG CTG CTC ACT CTT ATC ATT CTG TTG CCA GTA GTT TCA AAA TTT AGT TTT GTT 116
AGT CTC TCA GCA CCG CAG CAC TGG AGC TGT CCT GAA GGT ACT CTC GCA GGA AAT 170
GGG AAT TCT ACT TGT GTG GGT CCT GCA CCC TTC TTA ATT TTC TCC CAT GGA AAT 224
AGT ATC TTT AGG ATT GAC ACA GAA GGA ACC AAT TAT GAG CAA TTG GTG GTG GAT 278
GCT GGT GTC TCA GTG ATC ATG GAT TTT CAT TAT AAT GAG AAA AGA ATC TAT TGG 332
GTG GAT TTA GAA AGA CAA CTT TTG CAA AGA GTT TTT CTG AAT GGG TCA AGG CAA 386
GAG AGA GTA TGT AAT ATA GAG AAA AAT GTT TCT GGA ATG GCA ATA AAT TGG ATA 440
AAT GAA GAA GTT ATT TGG TCA AAT CAA CAG GAA GGA ATC ATT ACA GTA ACA GAT 494
ATG AAA GGA AAT AAT TCC CAC ATT CTT TTA AGT GCT TTA AAA TAT CCT GCA AAT 548
GTA GCA GTT GAT CCA GTA GAA AGG TTT ATA TTT TGG TCT TCA GAG GTG GCT GGA 602
AGC CTT TAT AGA GCA GAT CTC GAT GGT GTG GGA GTG AAG GCT CTG TTG GAG ACA 656
```


BNSDOCID: <WO_______9821348A1_I_>

TCA GAG AAA ATA ACA GCT GTG TCA TTG GAT GTG CTT GAT AAG CGG CTG TTT TGG 710
ATT CAG TAC AAC AGA GAA GGA AGC AAT TCT CTT ATT TGC TCC TGT GAT TAT GAT 764
GGA GGT TCT GTC CAC ATT AGT AAA CAT CCA ACA CAG CAT AAT TTG TTT GCA ATG 818
TCC CTT TTT GGT GAC CGT ATC TTC TAT TCA ACA TGG AAA ATG AAG ACA ATT TGG 872
ATA GCC AAC AAA CAC ACT GGA AAG GAC ATG GTT AGA ATT AAC CTC CAT TCA TCA 926
TTT GTA CCA CTT GGT GAA CTG AAA GTA GTG CAT CCA CTT GCA CAA CCC AAG GCA 980
GAA GAT GAC ACT TGG GAG CCT GAG CAG AAA CTT TGC AAA TTG AGG AAA GGA AAC 1034
TGC AGC AGC ACT GTG TGT GGG CAA GAC CTC CAG TCA CAC TTG TGC ATG TGT GCA 1088
GAG GGA TAC GCC CTA AGT CGA GAC CGG AAG TAC TGT GAA GAT GTT AAT GAA TGT 1142
GCT TTT TGG AAT CAT GGC TGT ACT CTT GGG TGT AAA AAC ACC CCT GGA TCC TAT 1196
TAC TGC ACG TGC CCT GTA GGA TTT GTT CTG CTT CCT GAT GGG AAA CGA TGT CAT 1250
CAA CTT GTT TCC TGT CCA CGC AAT GTG TCT GAA TGC AGC CAT GAC TGT GTT CTG 1304
ACA TCA GAA GGT CCC TTA TGT TTC TGT CCT GAA GGC TCA GTG CTT GAG AGA GAT 1358
GGG AAA ACA TGT AGC GGT TGT TCC TCA CCC GAT AAT GGT GGA TGT AGC CAG CTC 1412
TGC GTT CCT CTT AGC CCA GTA TCC TGG GAA TGT GAT TGC TTT CCT GGG TAT GAC 1466
CTA CAA CTG GAT GAA AAA AGC TGT GCA GCT TCA GGA CCA CAA CCA TTT TTG CTG 1520
TTT GCC AAT TCT CAA GAT ATT CGA CAC ATG CAT TTT GAT GGA ACA GAC TAT GGA 1574
ACT CTG CTC AGC CAG CAG ATG GGA ATG GTT TAT GCC CTA GAT CAT GAC CCT GTG 1628
GAA AAT AAG ATA TAC TTT GCC CAT ACA GCC CTG AAG TGG ATA GAG AGA GCT AAT 1682
ATG GAT GGT TCC CAG CGA GAA AGG CTT ATT GAG GAA GGA GTA GAT GTG CCA GAA 1736
GGT CTT GCT GTG GAC TGG ATT GGC CGT AGA TTC TAT TGG ACA GAC AGA GGG AAA 1790
TCT CTG ATT GGA AGG AGT GAT TTA AAT GGG AAA CGT TCC AAA ATA ATC ACT AAG 1844
GAG AAC ATC TCT CAA CCA CGA GGA ATT GCT GTT CAT CCA ATG GCC AAG AGA TTA 1898
TTC TGG ACT GAT ACA GGG ATT AAT CCA CGA ATT GAA AGT TCT TCC CTC CAA GGC 1952
CTT GGC CGT CTG GTT ATA GCC AGC TCT GAT CTA ATC TGG CCC AGT GGA ATA ACG 2006
ATT GAC TTC TTA ACT GAC AAG TTG TAC TGG TGC GAT GCC AAG CAG TCT GTG ATT 2060
GAA ATG GCC AAT CTG GAT GGT TCA AAA CGC CGA AGA CTT ACC CAG AAT GAT GTA 2114
GGT CAC CCA TTT GCT GTA GCA GTG TTT GAG GAT TAT GTG TGG TTC TCA GAT TGG 2168
GCT ATG CCA TCA GTA ATA AGA GTA AAC AAG AGG ACT GGC AAA GAT AGA GTA CGT 2222
CTC CAA GGC AGC ATG CTG AAG CCC TCA TCA CTG GTT GTG GTT CAT CCA TTG GCA 2276
AAA CCA GGA GCA GAT CCC TGC TTA TAT CAA AAC GGA GGC TGT GAA CAT ATT TGC 2330
AAA AAG AGG CTT GGA ACT GCT TGG TGT TCG TGT CGT GAA GGT TTT ATG AAA GCC 2384
TCA GAT GGG AAA ACG TGT CTG GCT CTG GAT GGT CAT CAG CTG TTG GCA GGT GGT 2438
GAA GTT GAT CTA AAG AAC CAA GTA ACA CCA TTG GAC ATC TTG TCC AAG ACT AGA 2492
GTG TCA GAA GAT AAC ATT ACA GAA TCT CAA CAC ATG CTA GTG GCT GAA ATC ATG 2546
GTG TCA GAT CAA GAT GAC TGT GCT CCT GTG GGA TGC AGC ATG TAT GCT CGG TGT 2600
ATT TCA GAG GGA GAG GAT GCC ACA TGT CAG TGT TTG AAA GGA TTT GCT GGG GAT 2654
GGA AAA CTA TGT TCT GAT ATA GAT GAA TGT GAG ATG GGT GTC CCA GTG TGC CCC 2708
CCT GCC TCC TCC AAG TGC ATC AAC ACC GAA GGT GGT TAT GTC TGC CGG TGC TCA 2762
GAA GGC TAC CAA GGA GAT GGG ATT CAC TGT CTT GAT ATT GAT GAG TGC CAA CTG 2816
GGG GTG CAC AGC TGT GGA GAG AAT GCC AGC TGC ACA AAT ACA GAG GGA GGC TAT 2870
ACC TGC ATG TGT GCT GGA CGC CTG TCT GAA CCA GGA AAT AGT GAC TCT GAA TGT 2924
CCC CTG TCC CAC GAT GGG TAC TGC CTC CAT GAT GGT GTG TGC ATG TAT ATT GAA 2978
GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT GTT GGC TAC ATC GGG GAG CGA TGT 3032
CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC CTG ATT TGC CCT GAC TCT ACT 3086
CCA CCC CCT CAC CTC AGG GAA GAT GAC CAC CAC TAT TCC GTA AGA CAC GCT GGC 3140
CAC GGG CAG CAG CAG AAG GTC ATC GTG GTG GCT GTC TGC GTG GTG GTG CTT GTC 3194
ATG CTG CTC CTC CTG AGC CTG TGG GGG GCC CAC TAC TAC AGG ACT CAG AAG CTG 3248
CTA TCG AAA AAC CCA AAG AAT CCT TAT GAG GAG TCG AGC AGA GAT GTG AGG AGT 3302
CGC AGG CCT GCT GAC ACT GAG GAT GGG ATG TCC TCT TGC CCT CAA CCT TGG TTT 3356
GTG GTT ATA AAA GAA CAC CAA GAC CTC AAG AAT GGG GGT CAA CCA GTG GCT GGT 3410
GAG GAT GGC CAG GCA GCA GAT GGG TCA ATG CAA CCA ACT TCA TGG AGG CAG GAG 3464
CCC CAG TTA TGT GGA ATG GGC ACA GAG CAA GGC TGC TGG ATT CCA GTA TCC AGT 3518
GAT AAG GGC TCC TGT CCC CAG GTA ATG GAG CGA AGC TTT CAT ATG CCC TCC TAT 3572
GGG ACA CAG ACC CTT GAA GGG GGT GTC GAG AAG CCC CAT TCT CTC CTA TCA GCT 3626
AAC CCA TTA TGG CAA CAA AGG GCC CTG GAC CCA CCA CAC CAA ATG GAG CTG ACT 3680
CAG TGA 3686

AAACTGGAAT TAAAAGGAAA GTCAAGAAGA ATGAACTATG TCGATGCACA GTATCTTTTC 3746
TTTCAAAAGT AGAGCAAAAC TATAGGTTTT GGTTCCACAA TCTCTACGAC TAATCACCTA 3806
CTCAATGCCT GGAGACAGAT ACGTAGTTGT GCTTTTGTTT GCTCTTTTAA GCAGTCTCAC 3866
TGCAGTCTTA TTTCCAAGTA AGAGTACTGG GAGAATCACT AGGTAACTTA TTAGAAACCC 3926
AAATTGGGAC AACAGTGCTT TGTAAATTGT GTTGTCTTCA GCAGTCAATA CAAATAGATT 3986
TTTGTTTTTG TTGTTCCTGC AGCCCCAGAA GAAATTAGGG GTTAAAGCAG ACAGTCACAC 4046
TGGTTTGGTC AGTTACAAAG TAATTTCTTT GATCTGGACA GAACATTTAT ATCAGTTTCA 4106
TGAAATGATT GGAATATTAC AATACCGTTA AGATACAGTG TAGGCATTTA ACTCCTCATT 4166
GGCGTGGTCC ATGCTGATGA TTTTGCCAAA ATGAGTTGTG ATGAATCAAT GAAAAATGTA 4226

```
ATTTAGAAAC TGATTTCTTC AGAATTAGAT GGCCTTATTT TTTAAAATAT TTGAATGAAA      4286
ACATTTTATT TTTAAAATAT TACACAGGAG GCCTTCGGAG TTTCTTAGTC ATTACTGTCC      4346
TTTTCCCCTA CAGAATTTTC CCTCTTGGTG TGATTGCACA GAATTTGTAT GTATTTTCAG      4406
TTACAAGATT GTAAGTAAAT TGCCTGATTT GTTTTCATTA TAGACAACGA TGAATTTCTT      4466
CTAATTATGA ATTC                                                        4480
```

(3) INFORMATION FOR SEQ ID NO: 2:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 783bp
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: double strands
(D) TOPOLOGY: unknown
(ii) MOLECULE TYPE:
(A) DESCRIPTION: sense orientation of five copies of mature EGF concatemers
(iii) HYPOTHETICAL:
(iv) ANTI-SENSE: 5'-CGC GTC AAG GGT ... TCT CAG TGA TAA-3 end
(v) FRAGMENT TYPE: 4.5kb Smal/Hindlll double strands DNA fragment
(vi) ORIGINAL SOURCE:
(A) ORGANISM: kidney
(B) STRAIN: human
(C) INDIVIDUAL ISOLATE: Z.Dai, et al.
(D) DEVELOPMENTAL STAGE: adult
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: fetal human liver library
(B) CLONE: lambda CH4A;lambdaEMBL4; lambda GM1416
(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT:
(B) MAP POSITION:
(C) UNITS:
(ix) FEATURE:
(A) NAME/KEY: Concatemer of mature EGF fragment without linker
(B) LOCATION:
(C) IDENTIFICATION METHOD: PCR cloning
(D) OTHER INFORMATION:
(x) PUBLICATION INFORMATION:
(A) AUTHORS:
(B) TITLE:
(C) JOURNAL:
(D) VOLUME:
(E) ISSUE:
(F) PAGES:
(G) DATE:
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO:
FROM ____(position) TO ____ (position)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

AAT AGT GAC TCT GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC CTC CAT GAT GGT 54
GTG TGC ATG TAT ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT GTT GGC 108
TAC ATC GGG GAG CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC AAT 162
AGT GAC TCT GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC CTC CAT GAT GGT GTG 216
TGC ATG TAT ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT GTT GGC TAC 270
ATC GGG GAG CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC AAT AGT 324
GAC TCT GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC CTC CAT GAT GGT GTG TGC 378
ATG TAT ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT GTT GGC TAC ATC 432
GGG GAG CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC AAT AGT GAC 486

BNSDOCID: <WO_______9821348A1_I_>

```
TCT GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC CTC CAT GAT GGT GTG TGC ATG 540
TAT ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT GTT GGC TAC ATC GGG 594
GAG CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC AAT AGT GAC TCT 648
GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC CTC CAT GAT GGT GTG TGC ATG TAT 702
ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT GTT GGC TAC ATC GGG GAG 756
CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC                     795
```

(4) INFORMATION FOR SEQ ID NO: 3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 891bp
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: double strands
(D) TOPOLOGY: unknown
(ii) MOLECULE TYPE:
(A) DESCRIPTION: sense orientation concatemer of mature EGF fragments with linkers
(iii) HYPOTHETICAL:
(iv) ANTI-SENSE: 5'-GTC CAG AGC ... CAG TGA TAA-3 end
(v) FRAGMENT TYPE: 5-copies of 159bp concatemer mature EGF linked with linkers
(vi) ORIGINAL SOURCE:
(A) ORGANISM: kidney
(B) STRAIN: human
(C) INDIVIDUAL ISOLATE: Z. Dai, et al
(D) DEVELOPMENTAL STAGE: adult
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:
(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE:
(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT:
(B) MAP POSITION:
(C) UNITS:
(ix) FEATURE:
(A) NAME/KEY: concatemer of mature EGF linked with linkers
(B) LOCATION:
(C) IDENTIFICATION METHOD:PCR cloning
(D) OTHER INFORMATION: Cleavage sites at 142-165, 307-331, 465-489, 631-655.
(x) PUBLICATION INFORMATION:
(A) AUTHORS:
(B) TITLE:
(C) JOURNAL:
(D) VOLUME:
(E) ISSUE:
(F) PAGES:
(G) DATE:
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO:
FROM ____ (position) TO ____ (position)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
AAT AGT GAC TCT GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC CTC CAT GAT GGT 54
GTG TGC ATG TAT ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT GTT GGC 108
TAC ATC GGG GAG CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC GGC 162
GGA AGA GTT AAC TGC ATG CAG AAT AGT GAC TCT GAA TGT CCC CTG TCC CAC GAT 216
GGG TAC TGC CTC CAT GAT GGT GTG TGC ATG TAT ATT GAA GCA TTG GAC AAG TAT 270
GCA TGC AAC TGT GTT GTT GGC TAC ATC GGG GAG CGA TGT CAG TAC CGA GAC CTG 324
AAG TGG TGG GAA CTG CGC GGC GGA AGA GTT AAC TGC ATG CAG AAT AGT GAC TCT 378
GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC CTC CAT GAT GGT GTG TGC ATG TAT 432
ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT GTT GTT GGC TAC ATC GGG GAG 486
CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA CTG CGC GGC GGA AGA GTT AAC 540
TGC ATG CAG AAT AGT GAC TCT GAA TGT CCC CTG TCC CAC GAT GGG TAC TGC CTC 594
CAT GAT GGT GTG TGC ATG TAT ATT GAA GCA TTG GAC AAG TAT GCA TGC AAC TGT 648
GTT GTT GGC TAC ATC GGG GAG CGA TGT CAG TAC CGA GAC CTG AAG TGG TGG GAA 702
CTG CGC GGC GGA AGA GTT AAC TGC ATG CAG AAT AGT GAC TCT GAA TGT CCC CTG 756
TCC CAC GAT GGG TAC TGC CTC CAT GAT GGT GTG TGC ATG TAT ATT GAA GCA TTG 810
GAC AAG TAT GCA TGC AAC TGT GTT GTT GGC TAC ATC GGG GAG CGA TGT CAG TAC 864
CGA GAC CTG AAG TGG TGG GAA CTG CGC                                     891
```


BNSDOCID: <WO________9821348A1_I_>

(5) INFORMATION FOR SEQ ID NO: 4:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 330bp
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double strands
(D) TOPOLOGY: unknown
(ii) MOLECULE TYPE:
(A) DESCRIPTION: upstream enhancer (from -343 to -90 bp) of 35S promoter
(iii) HYPOTHETICAL:
(iv) ANTI-SENSE:
(v) FRAGMENT TYPE: 253bp upstream of 35S promoter enhancer element
(vi) ORIGINAL SOURCE:
(A) ORGANISM: cauliflower mosaic virus (CaMV)
(B) STRAIN: Cabb B-D
(C) INDIVIDUAL ISOLATE: Z.Dai, et al
(D) DEVELOPMENTAL STAGE:
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:

(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE:
(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT:
(B) MAP POSITION:
(C) UNITS:
(ix) FEATURE:
(A) NAME/KEY: 35S promoter B-domain enhancer
(B) LOCATION:
(C) IDENTIFICATION METHOD: standard cloning
(D) OTHER INFORMATION: B-domain of 35S promoter from EcoR V site to Hind II site (upstream enhancer region from -343 to -90 bp)
(x) PUBLICATION INFORMATION:
(A) AUTHORS:
(B) TITLE:
(C) JOURNAL:
(D) VOLUME:
(E) ISSUE:
(F) PAGES:
(G) DATE:
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO:
FROM ____ (position) TO ____ (position)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
GTCAACATGG TGGAGCACGA CACACTTGTC TACTCCAAAA ATATCAAAGA TACAGTCTCA 60
GAAGACCAAA GGGCAATTGA GACTTTTCAA CAAAGGGTAA TATCCGGAAA CCTCCTCGGA 120
TTCCATTGCC CAGCTATCTG TCACTTTATT GTGAAGATAG TGGAAAAGGA AGGTGGCTCC 180
TACAAATGCC ATCATTGCGA TAAAGGAAAG GCCATCGTTG AAGATGCCTC TGCCGACAGT 240
GGTCCCAAAG ATGGACCCCC ACCCACGAGG AGCATCGTGG AAAAAGAAGA CGTTCCAACC 300
ACGTCTTCAA AGCAAGTGGA TTGATGTGAT                                  330
```

(6) INFORMATION FOR SEQ ID NO: 5:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1441bp
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: double strands
(D) TOPOLOGY: unknown
(ii) MOLECULE TYPE:
(A) DESCRIPTION: 5'-untranscription region of chl a/b binding protein
(iii) HYPOTHETICAL:
(iv) ANTI-SENSE:
(v) FRAGMENT TYPE: 11kb EcoR 1 fragment
(vi) ORIGINAL SOURCE:
(A) ORGANISM: whole plants
(B) STRAIN: Arabidopsis
(C) INDIVIDUAL ISOLATE: Ha et al
(D) DEVELOPMENTAL STAGE: 30 day old seedlings
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: genomic DNA library
(B) CLONE: lambda bAT1005
(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT:
(B) MAP POSITION:
(C) UNITS:
(ix) FEATURE:
(A) NAME/KEY: arabidopsis cab1 gene promoter
(B) LOCATION:
(C) IDENTIFICATION METHOD: cross-hybridization
(D) OTHER INFORMATION:
(x) PUBLICATION INFORMATION:
(A) AUTHORS:
(B) TITLE:
(C) JOURNAL:
(D) VOLUME:
(E) ISSUE:
(F) PAGES:
(G) DATE:
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO:
FROM ____(position) TO ____ (position)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
GAATTCATCA ACAAATTACT CCTCAATCAC ACTCCTATAG AAAACGGTTT AAGCTATCAT 60
TACATGTCTA GTTGGTTTTA CTCAGCCCTA GAAGTGTTGT TTATTGCATC ACTTTCCACG 120
AAGCACAATT TTTCTTTTTT ACAATCACTA GACCTCACAG GCTCACACAT ATGCTTTAGA 180
GCACATTCTA AACTTTGAAC TATAAAAGCT GTTAACACTA ATACACTATG CGTTCTTTTT 240
TGCTCCAAAC ACTTTTGATC CATTATTAGG AGACACTCCA CTTAGAAAGA TTTTCTAATC 300
CTTTGGTCAA CTAGGAAGTT CAAGGTTTTT CTAAACAGAA ATTCATTTCA CAAGTAATTT 360
AATTTATAAG GAAATGAATA GAGAAATCAA ATCATTGAAG AACTACAAAA TATAGATTCA 420
AGGTCAGGTC TAAGAAAATA TTCCTGAAGC TCAAAAAAGA GTTTTCCTCT CACATTATAG 480
AATTGGCCTT TACTTCAACA TTTTCCCACC TATTCCACAT TTGGTCAGAA CATTTTTAAT 540
TACTTGTGGA TCAATTTCCG GTTGAAATGG GTTTGGTGAA TATCCGGTTC AGTTATATGG 600
TGGCCGTTGG AATTGGCTTA TTAGTTGTGG CCGTTGTTGA AGCCGTTGGT ATTGGTAAGG 660
GAGAAGCAGA CTTGTGGCTA TGAGTCTATG ACCATGACTC GTGATTATGG AGCTGTCTTA 720
TGACCCTGAC CATCACCTTG ATCTGGTGGA TTCCAATGTT TTCTTCTTCT TCTAATAAAA 780
TATTATGGTC AATACAGGTG CTAATTAAGA TGGTAATAAT TTCTTATGTT TCTGTGGTAA 840
AGTTTGATTC AATTCCGTAG TTTTAGATAA TCTTATTTCC ATACATAAAT TTTATAGTTT 900
TATCTACTTT GTTCTTATGT TTTATCTCTA GCCAAGAGTT ATTATTATTA TCAGAAGAAG 960
AAAAAAAAAA GAAGCATATA TACAAAAGGT TTAATAAAAT GTATTATACA AGGCAATTAT 1020
```

```
CCAAATTTTT TTTGTTTTGG TTTACATTGA TGCTCTCAGG ATTTCATAAG GATAGAGAGA 1080
TCTATTCGTA TACGTGTCAC GTCATGAGTG GGTGTTTCGC CAATCCATGA AACGCACCTA 1140
GATATCTAAA ACACATATCA ATTGCGAATC TGCGAAGTGC GAGCCATTAA CCACGTAAGC 1200
AAACAAACAA TCTAAACCCC AAAAAAAATC TATGACTAGC CAATAGCAAC CTCAGAGATT 1260
GATATTTCAA GATAAGACAG TATTTAGATT TCTGTATTAT ATATAGCGAA AATCGCATCA 1320
ATACCAAACC ACCCATTTCT TGGCTTACAA CAACAAATCT TAAACGTTTT ACTTTGTGCT 1380
GCACTACTCA ACCTTAATGG CCGCCTCAAC AATGGCTCTC TCCTCCCCTG CCTTCGCCGG 1440
T                                                                 1441
```

(7) INFORMATION FOR SEQ ID NO: 6:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 832bp
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: double strands
(D) TOPOLOGY: unknown
(ii) MOLECULE TYPE:
(A) DESCRIPTION: CaMV 35S 5'-untranscription upstream
(iii) HYPOTHETICAL:
(iv) ANTI-SENSE:
(v) FRAGMENT TYPE: Alu 1 (from 7143bp)-EcoR1(to 7517bp)
(vi) ORIGINAL SOURCE:
(A) ORGANISM: cauliflower mosaic virus
(B) STRAIN: cM4-184
(C) INDIVIDUAL ISOLATE: RJ Shepherd
(D) DEVELOPMENTAL STAGE:
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: genomic library of CM4-184
(B) CLONE: pOS-1
(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT:
(B) MAP POSITION:
(C) UNITS:
(ix) FEATURE:
(A) NAME/KEY: CaMV 35S promoter
(B) LOCATION:
(C) IDENTIFICATION METHOD: cross-hybridization
(D) OTHER INFORMATION:
(x) PUBLICATION INFORMATION:
(A) AUTHORS:
(B) TITLE:
(C) JOURNAL:
(D) VOLUME:
(E) ISSUE:
(F) PAGES:
(G) DATE:
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO:
FROM ____(position) TO ____ (position)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
CCCACA GATGGTTAGA GAGGCTTACG CAGCAGGTCT CATCAAGACG ATCTACCCGA    56
GCAATAATCT CCAGGAAATC AAATACCTTC CCAAGAAGGT TAAAGATGCA GTCAAAAGAT 116
TCAGGACTAA CTGCATCAAG AACACAGAGA AAGATATATT TCTCAAGATC AGAAGTACTA 176
TTCCAGTATG GACGATTCAA GGCTTGCTTC ACAAACCAAG GCAAGTAATA GAGATTGGAG 236
TCTCTAAAAA GGTAGTTCCC ACTGAATCAA AGGCCATGGA GTCAAAGATT CAAATAGAGG 296
ACCTAACAGA ACTCGCCGTA AAGACTGGCG AACAGTTCAT ACAGAGTCTC TTACGACTCA 356
ATGACAAGAA GAAAATCTTC GTCAACATGG TGGAGCACGA CACACTTGTC TACTCCAAAA 416
ATATCAAAGA TACAGTCTCA GAAGACCAAA GGGCAATTGA GACTTTTCAA CAAAGGGTAA 476
TATCCGGAAA CCTCCTCGGA TTCCATTGCC CAGCTATCTG TCACTTTATT GTGAAGATAG 536
TGGAAAAGGA AGGTGGCTCC TACAAATGCC ATCATTGCGA TAAAGGAAAG GCCATCGTTG 596
AAGATGCCTC TGCCGACAGT GGTCCCAAAG ATGGACCCCC ACCCACGAGG AGCATCGTGG 656
AAAAAGAAGA CGTTCCAACC ACGTCTTCAA AGCAAGTGGA TTGATGTGAT ATCTCCACTG 716
ACGTAAGGGA TGACGCACAA TCCCACTATC CTTCGCAAGA CCCTTCCTCT ATATAAGGAA 776
GTTCATTTCA TTTGGAGAGA ACACGGGGGA CTCTAGAGGA TCCCCGGGTG GTCAGT     832
```

(8) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 473bp
   (B) TYPE: Nucleic acid
   (C) STRANDEDNESS: double strands
   (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE:
   (A) DESCRIPTION: 5'-untranscription upstream of ribosomal protein L34

(iii) HYPOTHETICAL:

(iv) ANTI-SENSE:

(v) FRAGMENT TYPE: 1500bp BamH-Hind 111

(vi) ORIGINAL SOURCE:
   (A) ORGANISM: tobacco NT1 cells
   (B) STRAIN:
   (C) INDIVIDUAL ISOLATE: Z.Dai, et al
   (D) DEVELOPMENTAL STAGE: 3 days old
   (E) HAPLOTYPE:
   (F) TISSUE TYPE:
   (G) CELL TYPE:
   (H) CELL LINE: NT1
   (I) ORGANELLE:

(vii) IMMEDIATE SOURCE:
   (A) LIBRARY: genomic library
   (B) CLONE: TSC 40

(viii) POSITION IN GENOME:
   (A) CHROMOSOME/SEGMENT:
   (B) MAP POSITION:
   (C) UNITS:

(ix) FEATURE:
   (A) NAME/KEY: RPL-34 promoter
   (B) LOCATION:
   (C) IDENTIFICATION METHOD: plaque hybridization
   (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:
   (A) AUTHORS:
   (B) TITLE:
   (C) JOURNAL:
   (D) VOLUME:
   (E) ISSUE:
   (F) PAGES:
   (G) DATE:
   (H) DOCUMENT NUMBER:
   (I) FILING DATE:
   (J) PUBLICATION DATE:
   (K) RELEVANT RESIDUES IN SEQ ID NO:
      FROM ____(position) TO ____ (position)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
AGATCTCT CTTTGTATTC TTATTGATGT ACTGGTTTGA AGATGAATAA AATCTTTCAT   58
TCCACCAAAA AAAGAATGAA AATAAAATTT TAATATACAT GTTGATATAG ACAAAGAAGA 118
AAAAAAAAGT TGTGATTACA TTTATTGACT ATTTGATGCC AATATCTATA ACTAGAGCTA 178
TTTTCTATCA ATTATATGGG TATGTTGTTA TACCATGCCA AAACCTCAAT TCATAATGTG 238
CTTGTTTAAA CCCAGTTTAA TGGGCTAACA TGTTGATGGG CTTATAGGCC CGTCTGATTT 298
CCTTGCCAGA CACTAGTAAG TAAATGATTC TATCATCCAA TATCAACCGT GGGATCTAGG 358
GCTTGTCCCA CTTATATACA CTACATATAT TTAACTTTCC TTTAGCCCTT CTGCTTCAGC 418
CCCCAAAACA AAGAAAGAAG CTACAGAGAG AATAGCAGCG CCGCCGTGAA AAATG      473
```

(9) INFORMATION FOR SEQ ID NO: 8:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1162bp
(B) TYPE: Nucleic acid
(C) STRANDEDNESS: double strands
(D) TOPOLOGY: unknown
(ii) MOLECULE TYPE:
(A) DESCRIPTION: 5'-untranscription region of 35S gene from CaMV with 2 copies of B domains
(iii) HYPOTHETICAL:
(iv) ANTI-SENSE:
(v) FRAGMENT TYPE: 253bp Hindlll/EcoRV fragment + 343bp Hind ll/EcoRl fragment
(vi) ORIGINAL SOURCE:
(A) ORGANISM: whole cell
(B) STRAIN: CM4-184
(C) INDIVIDUAL ISOLATE: Z.Dai, et al
(D) DEVELOPMENTAL STAGE:
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:
(vii) IMMEDIATE SOURCE:
(A) LIBRARY: genomic library of CM4-184
(B) CLONE: POS-1
(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT:
(B) MAP POSITION:
(C) UNITS:
(ix) FEATURE:
(A) NAME/KEY: CaMV 35S promoter with duplication of upstream B domain
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:
(x) PUBLICATION INFORMATION:
(A) AUTHORS:
(B) TITLE:
(C) JOURNAL:
(D) VOLUME:
(E) ISSUE:
(F) PAGES:
(G) DATE:
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO:
FROM ____ (position) TO ____ (position)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
CCC ACAGATGGTT AGAGAGGCTT ACGCAGCAGG TCTCATCAAG ACGATCTACC        53
CGAGCAATAA TCTCCAGGAA ATCAAATACC TTCCCAAGAA GGTTAAAGAT GCAGTCAAAA 113
GATTCAGGAC TAACTGCATC AAGAACACAG AGAAAGATAT ATTTCTCAAG ATCAGAAGTA 173
CTATTCCAGT ATGGACGATT CAAGGCTTGC TTCACAAACC AAGGCAAGTA ATAGAGATTG 233
GAGTCTCTAA AAAGGTAGTT CCCACTGAAT CAAAGGCCAT GGAGTCAAAG ATTCAAATAG 293
AGGACCTAAC AGAACTCGCC GTAAAGACTG GCGAACAGTT CATACAGAGT CTCTTACGAC 353
TCAATGACAA GAAGAAAATC TTCGTCAACA TGGTGGAGCA CGACACACTT GTCTACTCCA 413
AAAATATCAA AGATACAGTC TCAGAAGACC AAAGGGCAAT TGAGACTTTT CAACAAAGGG 473
TAATATCCGG AAACCTCCTC GGATTCCATT GCCCAGCTAT CTGTCACTTT ATTGTGAAGA 533
TAGTGGAAAA GGAAGGTGGC TCCTACAAAT GCCATCATTG CGATAAAGGA AAGGCCATCG 593
TTGAAGATGC CTCTGCCGAC AGTGGTCCCA AAGATGGACC CCCACCCACG AGGAGCATCG 653
TGGAAAAAGA AGACGTTCCA ACCACGTCTT CAAAGCAAGT GGATTGATGT GATAACATGG 713
TGGAGCACGA CACACTTGTC TACTCCAAAA ATATCAAAGA TACAGTCTCA GAAGACCAAA 773
GGGCAATTGA GACTTTTCAA CAAAGGGTAA TATCCGGAAA CCTCCTCGGA TTCCATTGCC 833
CAGCTATCTG TCACTTTATT GTGAAGATAG TGGAAAAGGA AGGTGGCTCC TACAAATGCC 893
ATCATTGCGA TAAAGGAAAG GCCATCGTTG AAGATGCCTC TGCCGACAGT GGTCCCAAAG 953
```


BNSDOCID: <WO________9821348A1_I_>

```
ATGGACCCCC ACCCACGAGG AGCATCGTGG AAAAAGAAGA CGTTCCAACC ACGTCTTCAA 1013
AGCAAGTGGA TTGATGTGAT ATCTCCACTG ACGTAAGGGA TGACGCACAA TCCCACTATC 1073
CTTCGCAAGA CCCTTCCTCT ATATAAGGAA GTTCATTTCA TTTGGAGAGA ACACGGGGGA 1133
CTCTAGAGGA TCCCCGGGTG GTCAGT                                      1159
```


BNSDOCID: <WO_________9821348A1_I_>

CLAIMS

We claim:

1. A method of producing human growth factors from plant cells, comprising the steps of:

(a) obtaining a positive transformant of the plant cells, the positive transformant carrying genetic material encoding the production of a human growth factor with a length of at least 200 amino acids;

(b) cultivating the positive transformant; and

(c) obtaining the human growth factors.

2. The method as recited in claim 1, wherein obtaining the positive transformant has the step of:

modifying a chimeric cDNA encoding the human growth factor with a length of at least 200 amino acids, for subcloning into a plant expression vector.

3. The method as recited in claim 2, further comprising the steps of:

(a) subcloning the chimeric cDNA into the plant expression vector and obtaining a subcloned plant expression vector;

(b) transferring the subcloned plant expression vector into a plurality of plant cells;

(c) selecting a plurality of positive transformants from the plurality of plant cells on an antibiotic selective media;

(d) permitting growth of the portion of the plurality of plant cells in whole plants or suspensions; and

(e) extracting a liquid containing the human growth factor from the plurality of transgenic plant cells.

4. The method as recited in claim 3, wherein transferring is by direct particle bombardment.

5. The method as recited in claim 3, wherein transferring is by Agrobacterium mediated transformation.

6. The method as recited in claim 5, wherein Agrobacterium mediated transformation comprises the steps of:

(a) placing the subcloned plant expression vector to an agrobacterium;

(b) co-cultivating the Agrobacterium containing the subcloned plant expression vector with the plurality of plant cells.

7. The method as recited in claim 1, wherein the step of cultivating is with a whole plant.

8. The method as recited in claim 1, wherein the step of cultivating is with a plant tissue culture.

9. The method as recited in claim 1, wherein the step of obtaining is selected from the group consisting of ultrafiltration, affinity chromatography, and electrophoresis.

10. The method as recited in claim 1, wherein the length of at least 200 amino acids is obtained by cloning a cDNA.

11. The method as recited in claim 10, wherein said cDNA is a pre-pro-EGF cDNA.

12. The method as recited in claim 1, wherein the length of at least 200 amino acids is obtained by synthesizing a cDNA.

13. The method as recited in claim 12, wherein said synthesizing is concatomerizing multiple gene copies to obtain the length of at least 200 amino acids.

14. The method as recited in claim 1, further comprising increasing an overall size of a gene to be expressed with a fusion construct encoding an hEGF linked to a protein that is efficiently produced in plant systems.

15. The method as recited in claim 1, wherein said human growth factor is selected from the group consisting of epidermal growth factor (EGF), transforming growth factor (TGF), vascular endothelial growth factor (VEGF), platelet-derived growth factor (PDGF), fibroblast growth factor (FGF), tumor necrosis factor (TNF), heparin-binding epidermal growth factor (HBEGF), insulin-like growth factor (ILGF), platelet-derived endothelial cell growth factor (PDECGF), platelet-derived angiogenesis factor (PDAF), and bone-and-cartilage inducing growth factor (BCIF) and combinations thereof.

16. The method as recited in claim 2, wherein modifying is by adding a regulatory element selected from the group consisting of leader sequences, signal peptides, transcription promoters or enhancers, and transcription terminators.

17. The method as recited in claim 2, wherein modifying a chimeric cDNA, comprises the steps of:

(a) adding said transcription promoter to the upstream or 5' end of the chimeric cDNA; and

(b) adding said transcription terminator to the downstream or 3' end of the chimeric cDNA.

18. The method as recited in claim 17, further comprising adding an additional regulatory element encoding a signal peptide, said additional regulatory

element added between the transcription promoter and the upstream 5' end of the chimeric cDNA.

19. The method as recited in claim 18, further comprising adding a regulatory element between the transcription promoter and the additional regulatory element encoding the signal peptide to enhance mRNA stability.

20. The method as recited in claim 18, further comprising adding a regulatory element at the downstream or 3' end of the chimeric cDNA to enhance mRNA stability.

21. The method as recited in claim 17, wherein transcription promoters limit growth factors production to a non-crop portion of a transgenic whole plant.

22. The method as recited in claim 21, wherein the transcription promoters are selected from the group consisting of an upstream enhancer region (-343 to -90 bp) of a CaMV 35S promoter, a chlorophyll a/b binding promoter (cab1) and combinations thereof.

23. The method as recited in claim 17, wherein the transcription promoters are selected from the group consisting of a modified 35S promoter, TSC29 promoter, TSC40 promoter and combinations thereof.

24. The method as recited in claim 23, wherein the modified 35S promoter is a 35S promoter modified by duplicating an upstream enhancer region (-343 to -90 bp) of the 35S promoter to increase transcription activity.

25. The method as recited in claim 2, wherein said cDNA is a pre-pro-EGF cDNA.

26. The method as recited in claim 25, wherein said pre-pro-EGF cDNA has approximately 4.5 kb, whereby overall titers of active hEGF in both whole plants and cell culture are increased.

27. The method as recited in claim 2, wherein the length of at least 200 amino acids is obtained by synthesizing the cDNA.

28. The method as recited in claim 27, wherein said synthesizing is concatomerizing multiple gene copies to obtain the length of at least 200 amino acids.

29. The method as recited in claim 28, wherein said multiple gene copies are an oligomeric polypeptide having of repeated hEGF domains.

30. The method as recited in claim 2, further comprising increasing an overall size of a gene to be expressed with a fusion construct encoding an hEGF linked to a protein that is efficiently produced in plant systems.

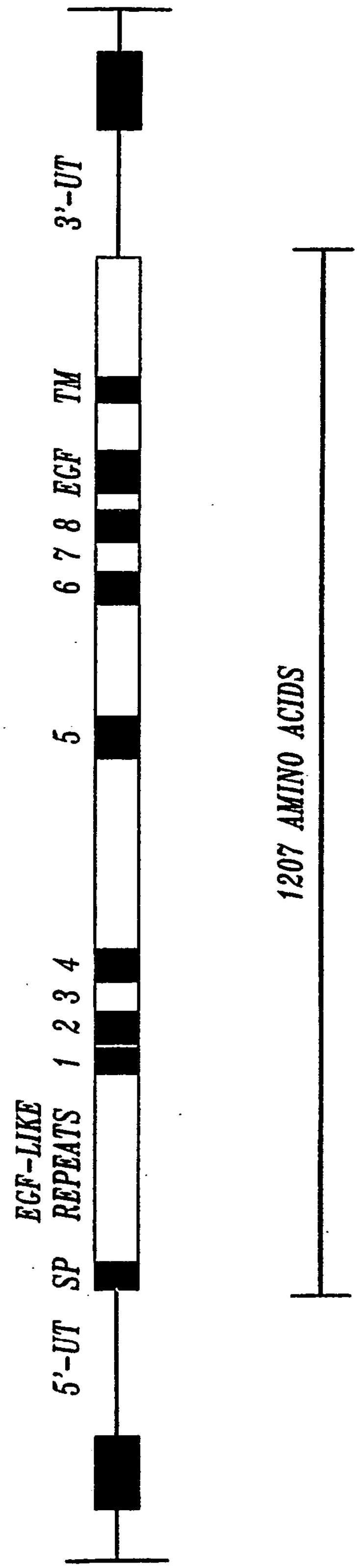

Fig. 1

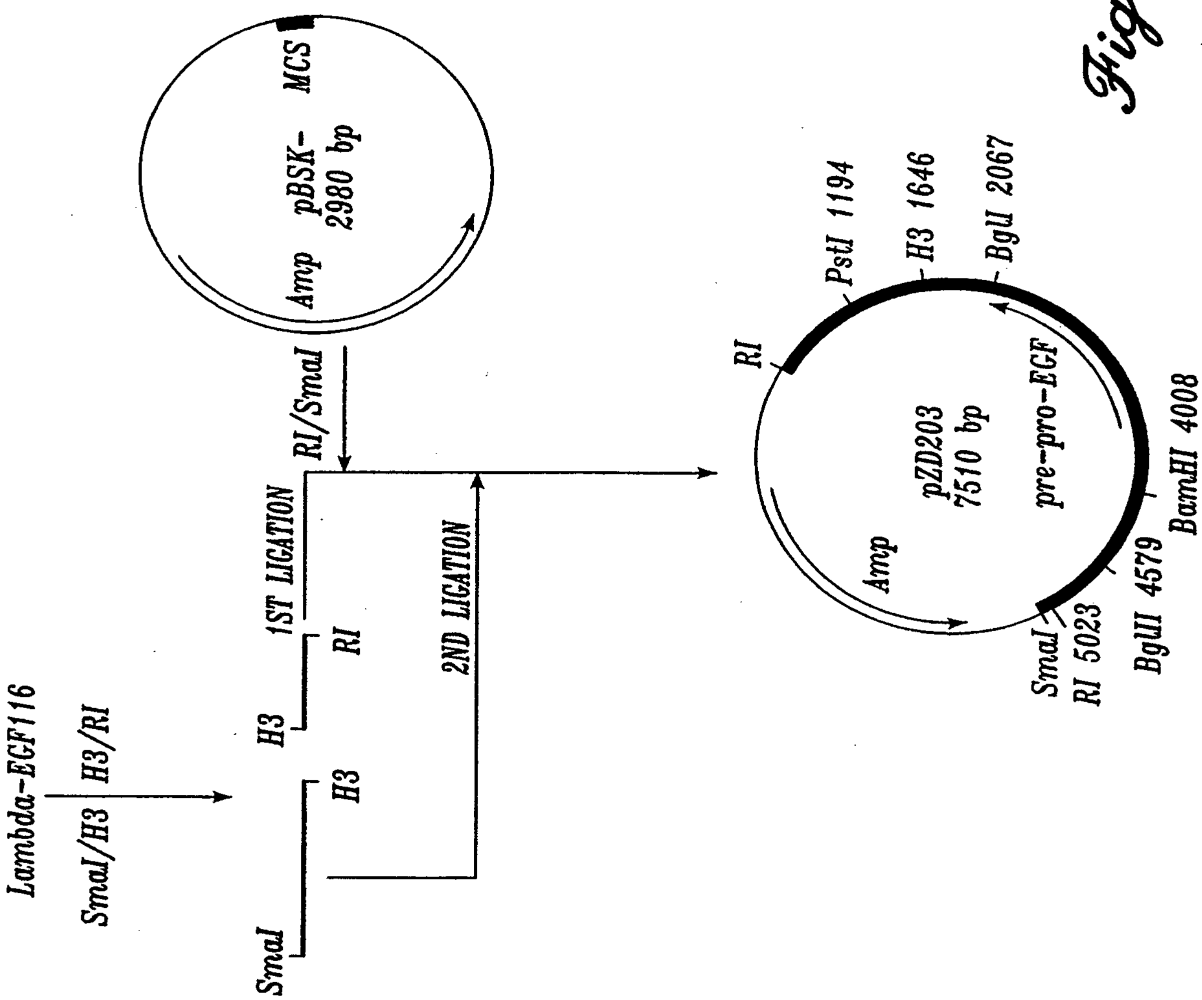
Lambda-EGF116
SmaI/H3
H3/RI
SmaI
H3
H3
RI
1ST LIGATION
2ND LIGATION
RI/SmaI
pBSK-
2980 bp
MCS
Amp
pZD203
7510 bp
pre-pro-EGF
Amp
RI
PstI 1194
H3 1646
BgII 2067
BamHI 4008
BgIII 4579
SmaI
RI 5023

Fig. 2

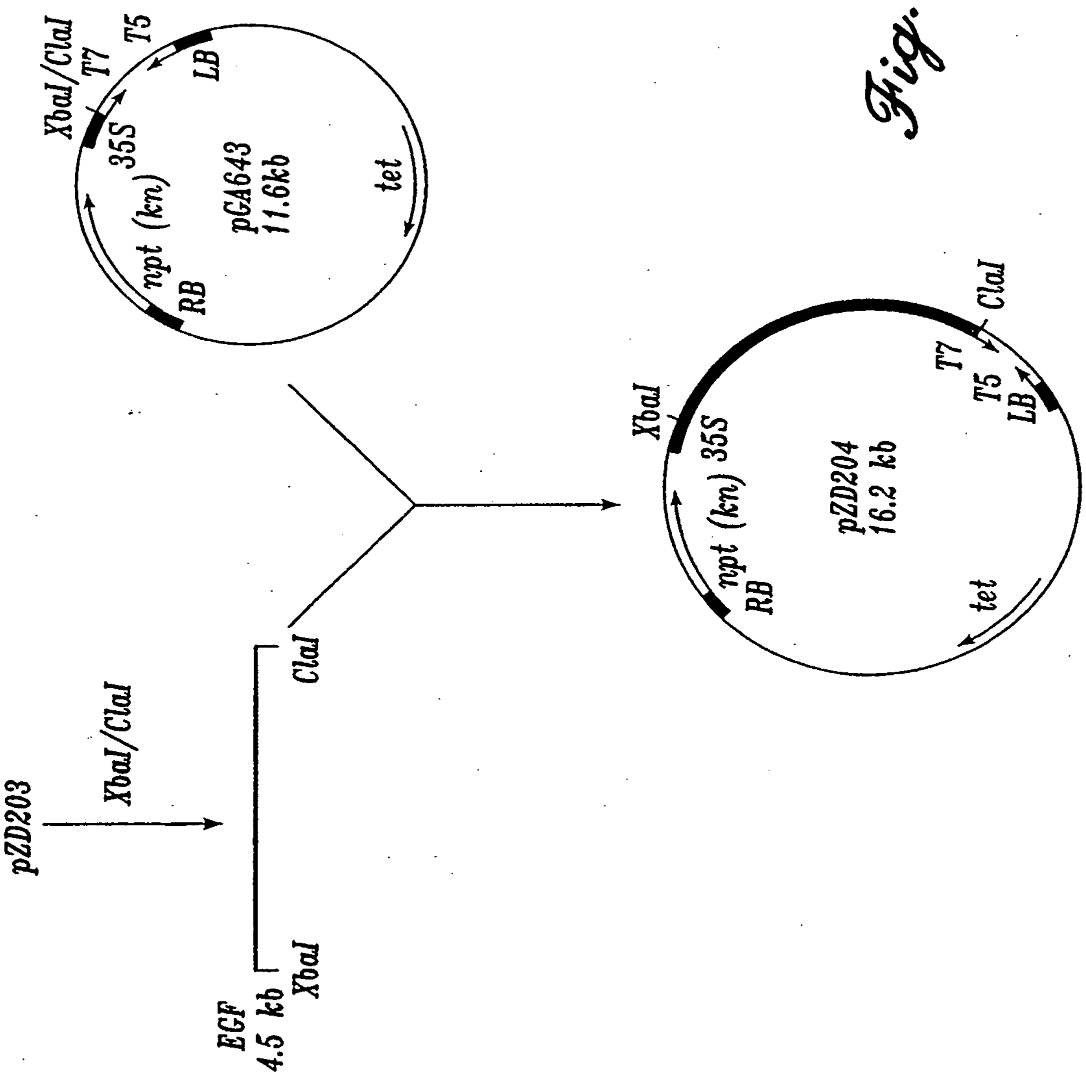
XbaI/ClaI
T7
T5
LB
35S
pCA643
11.6kb
npt (kn)
RB
tet
XbaI
ClaI
T7
T5
LB
35S
pZD204
16.2 kb
npt (kn)
RB
tet
pZD203
XbaI/ClaI
ClaI
XbaI
EGF
4.5 kb

Fig. 3

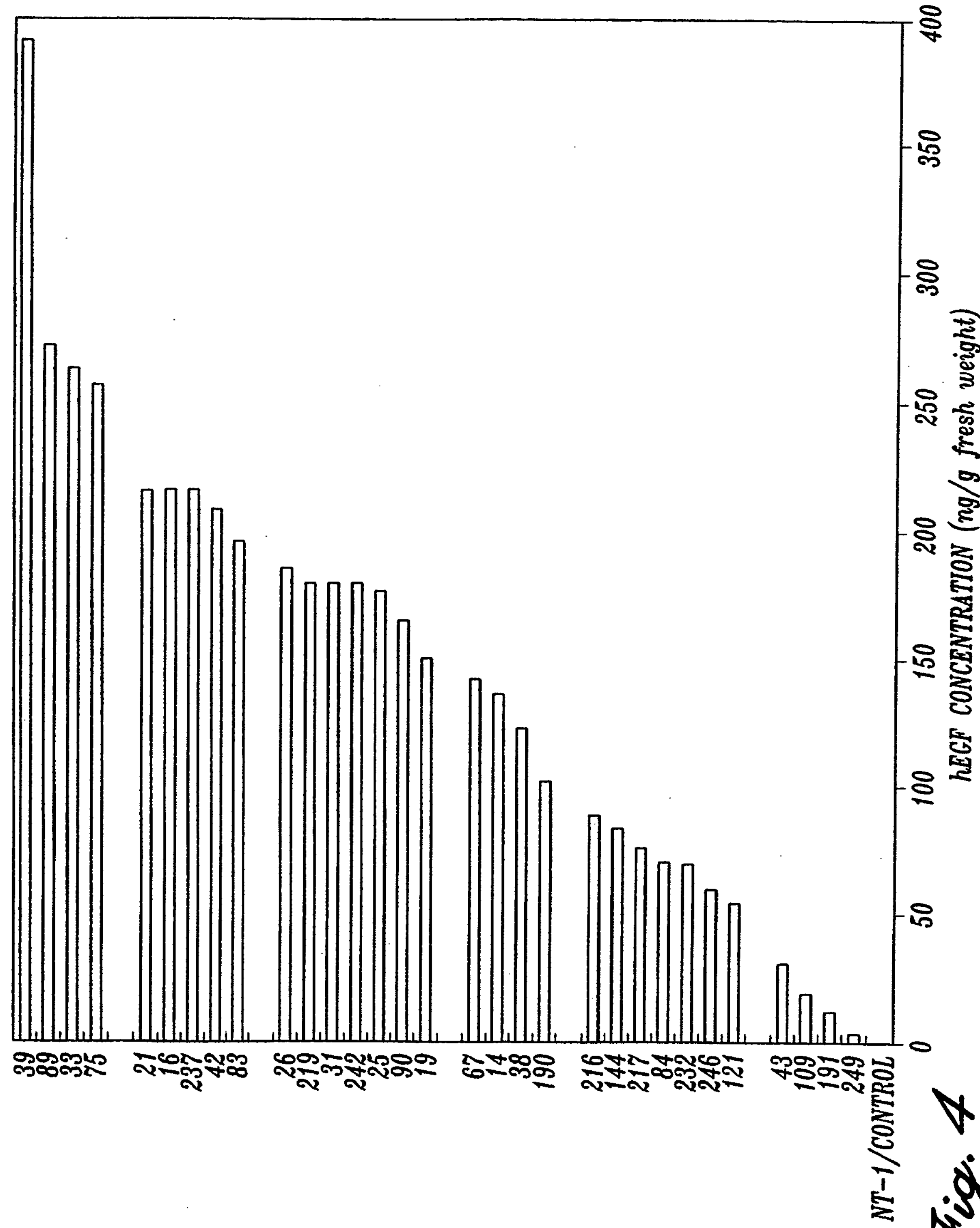
hEGF CONCENTRATION (ng/g fresh weight)
0
50
100
150
200
250
300
350
400
39
89
33
75
21
16
237
42
83
26
219
31
242
25
90
19
67
14
38
190
216
144
217
84
232
246
121
43
109
191
249
NT-1/CONTROL

Fig. 4

INTERNATIONAL SEARCH REPORT

Inter. ›nal Application No
PCT/US 97/20603

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 C12N15/12 C12N15/62 C12N5/10 C07K14/485
A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 87 00865 A (CALGENE INC) 12 February 1987 especially pages 5 + 6 see the whole document ---	1-3,5,6, 8,10,11, 15-18,25
A	MATSUMOTO S ET AL: "EXPRESSION OF HUMAN ERYTHROPOIETIN IN CULTURED TOBACCO CELLS" BIOSCIENCE BIOTECHNOLOGY BIOCHEMISTRY, vol. 57, no. 8, 1 August 1993, pages 1249-1252, XP000572270 especially page 1251, 1252 see the whole document ---	1-30

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
6 April 1998	21/04/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Holtorf, S

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	SALMANIAN, A-H., ET AL.: "synthesis and expression of the gene for human epidermal growth factor in transgenic potato plants" BIOTECHNOLOGY LETTERS, vol. 18, no. 9, 1996, pages 195-1098, XP002061490 especially page 1095 see the whole document	1-30
A	HIGO K -I ET AL: "EXPRESSION OF A CHEMICALLY SYNTHESIZED GENE FOR HUMAN EPIDERMAL GROWTH FACTOR UNDER THE CONTROL OF CAULIFLOWER MOSAIC VIRUS 35S PROMOTER IN TRANSGENIC TOBACCO" BIOSCIENCE BIOTECHNOLOGY BIOCHEMISTRY, vol. 57, no. 9, 1 September 1993, pages 1477-1481, XP000572271 cited in the application	1-30
A	WO 89 03887 A (PLANT GENETIC SYSTEMS NV) 5 May 1989 pages 3/4, page 10, line 25-36; page 48,49	1-30
A	WO 95 14099 A (RODRIGUEZ RAYMOND L ;UNIV CALIFORNIA (US)) 26 May 1995 see page 8	1-30
A	WO 93 21320 A (UNIV TECHNOLOGIES INT ;MOLONEY MAURICE M (CA)) 28 October 1993 pages 1,5,8-9; page 22-29; example 2	1-30

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter onal Application No
PCT/US 97/20603

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 8700865 A	12-02-87	US 4956282 A	11-09-90
		AU 603063 B	08-11-90
		AU 6196686 A	05-03-87
		DE 3687705 A	18-03-93
		EP 0233915 A	02-09-87
		JP 63500425 T	18-02-88
		US 5550038 A	27-08-96
		US 5629175 A	13-05-97
WO 8903887 A	05-05-89	CA 1337048 A	19-09-95
		AU 2811889 A	23-05-89
		AU 634987 B	11-03-93
		AU 4495189 A	01-05-90
		CA 2000661 A	14-04-90
		WO 9004032 A	19-04-90
		EP 0319353 A	07-06-89
		EP 0318341 A	31-05-89
		JP 2501802 T	21-06-90
		JP 3502644 T	20-06-91
		US 5487991 A	30-01-96
		AT 143694 T	15-10-96
		AT 140959 T	15-08-96
		DE 3855455 D	05-09-96
		DE 3855455 T	09-01-97
		DE 3855591 D	07-11-96
		DE 3855591 T	03-04-97
		EP 0723019 A	24-07-96
		US 5623067 A	22-04-97
WO 9514099 A	26-05-95	US 5693506 A	02-12-97
		AU 1289295 A	06-06-95
		CA 2176834 A	26-05-95
		EP 0788550 A	13-08-97
		JP 9509565 T	30-09-97
WO 9321320 A	28-10-93	AU 1583292 A	18-11-93
		AU 678154 B	22-05-97
		AU 3884493 A	08-11-93
		EP 0636179 A	01-02-95
		EP 0633940 A	18-01-95

Form PCT/ISA/210 (patent family annex) (July 1992)

BNSDOCID: <WO________9821348A1_I_>

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte .onal Application No
PCT/US 97/20603

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9321320 A		FI 944550 A	30-11-94
		FI 944841 A	14-12-94
		JP 7507199 T	10-08-95
		JP 7505525 T	22-06-95
		NO 943664 A	15-11-94
		NO 943914 A	09-12-94
		WO 9320216 A	14-10-93
		US 5650554 A	22-07-97

Form PCT/ISA/210 (patent family annex) (July 1992)

BNSDOCID: <WO_________9821348A1_I_>